Modül 2- ÖÇME- DEĞERLENDİRME

- Eğitimin gerçekleştirme istediği hedefler:

1- Uzak Hedefler: Devletin hedefleridir.
2- Genel Hedefler: Kurumların hedefleridir.
3- Özel Hedefler: Ders hedefleri, kazanımlardır.

- Eğitimin uzak hedefi, Atatürk'ün ortaya koyduğu, "Muhasır medeniyetler seviyesine çıkmak" tır.
- Eğitimin genel hedefi, toplumun bireylerden beklediği görev ve sorumluluklar doğrultusunda bilgi ve donanıma sahip olmaktır.

Eğitimin Ögeleri

GİRDİ	SÜREÇ	ÇIKTI
Eğitim için gerekli olan, Öğretmen Öğrenci Yönetim Çalışanlar Program	Eğitim durumları	Süreç sonunda öğrencilerin sahip olduğu nitelikler

ÖLÇME, ÖLÇÜT VE DEĞERLENDİRME

- Ölçme: Öğrencinin belli bir özelliğini gözleyerek o özelliğe sayı, sembol sıfat/kategori verme işlemidir. Nicelik bildirir. Funda sınavda 450 puan aldı, gibi.
- Ölçme Konuları:
  - → Akademik başarı
  - → Yetenek
  - → Özel gereksinim alanları
  - → İlgi
  - → Tutumlar
- Ölçüm: Ölçme sonucudur.
- Ölçüt: Ölçülen özellik hakkında karar alabilmek için dayanak noktası ya da referans aralığıdır.
- Değerlendirme: Bir ölçme sonucunu (ölçüm) en az bir ölçüte vurarak ölçülen nitelik hakkında karar verme işlemidir. Nitelik bildirir, Funda sınıfı geçti, gibi.

Değerlendirme Basamakları

- Ölçme- ölçüm
- Ölçüt
- Karar

ÖLÇME TÜRLERİ

TÜR	Özelliği	Örnek
Doğrudan(Temel) Ölçme	Gözleme dayalıdır. Beş duyu organıyla algılanır. Yargı barındırır.	Cinsiyet belirleme, boy ölçme, doğru okuma, diksiyon vb.
Dolaylı (Göstergeyle) Ölçme	Davranışlar aracılığıyla göstergeye dayalı ölçmedir.	Vergi vermenin iyi bir yurttaş olmanın göstergesi olması vb.
Türetilmiş Ölçme	Ölçülmek istenen özelliğin matematiksel bağıntıyla ölçülmesidir.	Zekâ yaşı hesaplamaları, Hız= yol/zaman vb.

ÖLÇÜT TÜRLERİ

Mutlak Ölçüt

- Grubun özelliklerine bakılmaz.
- Ölçme işlemi öncesinde ölçüt bellidir.
- Paydaşların ölçümleri birbirini etkilemez.

Ehliyet sınavı baraj puanı,
ALES notu
Geçme notunun 60 olması gibi kriterler

Bağıl Ölçüt

- Norm referanslı ölçüttür.
- Grubun belli bir özelliği belirlenir.
- Ölçüt ölçme sonrası belirlenebilir.
- Bireylerin ölçümleri birbirini etkileyebilir.

Aritmetik ortalama, standart sapma, moda denk gibi kriterler

Mutlak Değerlendirme: Mutlak ölçüt kullanılarak yapılan, öğrenci ile ilgili hayati kararlar alınırken uygulanır. Muafiyet sınavları, sınıf geçme vb.
Bağıl Değerlendirme: Bağıl ölçüt kullanılarak başvuran kişi sayısının çok, alınan lişi sayısının az olduğu sınavlarda, atama durumlarında uygulanır. KPSS, TUS; LGS vb. YKS gibi sınavlarda mutlak ve bağıl değerlendirme birlikte kullanılabilir.

AMACINA GÖRE DEĞERLENDİRME TÜRLERİ

Akademik Olanlar

1- Tanıma ve yerleştirmeye yönelik (diyagnostik)
- Amaç not vermek değildir.
- Okul/ kurum geneli seviye grupları oluşturulurken, hazırbulunuşluk belirlenirke, muafiyet durumlarında kullanılır.
- Zaman ve emek tasarrufu sağlar.

2- Biçimlendirme ve yetiştirmeye yönelik (formatif)
- Amaç not vermek değildir.
- İzleme, tarama, ünite sonu testleri gibi araçlarla kazanımın elde edilme düzeyini belirlemeye yöneliktir.

3- Değer biçmeye/ düzey belirlemeye yönelik
- Amaç not vermektir.
- Bitirme sınavları, yazılılar, sertifika sınavları gibi süreç sonu durum tespiti yapmaya yöneliktir.

Akademik Olmayanlar

4- Rehberlik amaçlı değerlendirmeler
- Özel eğitim ve mesleki eğitimde uygulanan değerlendirmeler

5- Program değerlendirme
- Programla ilgili yargıya ulaşmak için yapılır.

Bir Ölçme Aracında Bulunması Gereken Psikometrik Nitelikler

- ✓ Geçerlilik
- ✓ Güvenirlik
- ✓ Kullanışlılık

HATA TÜRLERİ

Bir ölçmede gözlenen aksaklıkların tamamı hata olarak adlandırılır.

Sabit Hata: Her ölçme işleminde aynı miktardaki hatadır. (Terazi ölçümü gibi)
Sistematik Hata: Ölçülen büyüklüğe öğretmene, koşullara göre oluşan hata.
Tesadüfî Hata: Şansla ortaya çıkan sonuca etki eden rastlantısal hatadır.

Tesadüfi Hatanın Kaynakları

- Öğretmen kaynaklı
- Ölçme aracı kaynaklı
- Birey kaynaklı
- Fiziksel ortam kaynaklı

KORELASYON

En az iki değişken arasında bir ilişki bulunup bulunmadığı, ilişki varsa yönü ve miktarını belirleyen tekniktir. r ile sembolize edilir.

✓ Pozitif Korelasyon: Değişkenler arasındaki orantılı ilişkidir. Öğretmenin ders saati arttıkça iş yükü artar, örneğindeki gibi.

✓ Negatif Korelasyon: İki değişken arasındaki ters orantılı ilişkidir. Ders süresi arttıkça dikkat azalır, örneğindeki gibi.

✓ Sıfır Korelasyon: Değişkenler arasındaki ilişkisizliktir. Öğretmenin boyu arttıkça başarı oranı artar, örneğindeki gibi.

GEÇERLİLİK SORGULAMA YÖNTEMLERİ

Kapsam Geçerliliği	Ölçülmek istenilen davranışların kapsamının mantıksal ya da istatistiksel yöntemlerle ölçülmesidir.	Belirtke tabloları, başarı testleri gibi
Ölçüt Dayanıklı Geçerlilik	Ölçme aracından elde edilen puanların ölçüt puanla karşılaştırıl-masıdır.	Sınavdan 50 ve üzeri alanlar geçer, gibi
Yordama Geçerliliği	Eldeki geçerli ve güvenilir verilere bakılarak henüz gerçekleşmemiş bir olguya dair tahmindir.	Seçme ve yönlendirme amaçlı testler (KPSS, LGS gibi testler. Bunlarda yordama geçerliliği yüksek olmalıdır.), yetenek testleri, mesleki kişilik envanteri gibi.
Zamandaş Geçerlilik	Daha önce geliştirilmiş ve geçerliği yüksek araçlardan alınan ölçüt puan ile yeni ölçme arasındaki ilişkidir.	Ölçüte bağlı olarak geçerlilik katsayısı ∓1'e yaklaştıkça artar, 0'a yaklaştıkça düşer.
Yapı Geçerliliği	Birbirleriyle ilgili olduğu düşünülen belli ögelerin ya da ögeler arasındaki ilişkilerin oluşturduğu bir örüntüdür. Bu anlamda, bir testin yapısını geçerleme süreci, temelde testin maddelerine verilen yanıtlar arasındaki ilişkilerin analizine dayanır.	Beşerî bilimlerde atılganlık, güvensizlik, içe dönüklük vb. özelliklerin ölçülmesi amacıyla

GÜVENİRLİLİK

Güvenilir bir ölçme aracı aynı özellikle ilgili olarak yapılan ölçmelerde yaklaşık aynı sonucu verir. İki ölçüm arasındaki uyumsuzluklar için gerçek puan ve değişkenler incelenmelidir. Hata miktarı arttıkça gözlenen puan ile gerçek puan arasındaki fark artar. Güvenirliliğin artması geçerliliği artırır.

GÜVENİRLİLİK YÖNTEMLERİ

YÖNTEM	UYGULAMASI	ÖRNEK
Test-Tekrar Test	Belli aralıklarla aynı testin aynı gruba yeniden yapılmasıdır. Bu yöntem başarı testleri için önerilmez. Yaş değişimi, zaman aralığı sorun oluşturur. Ölçülen özellik kararlı olmalıdır.	Zihin yetenekleri, kişilik testleri, ilgi envanterleri, tutum ölçekleri
Test Yarılama	Test iki eş parçaya bölünerek uygulanır. Bireylerin iki yarıdan aldıkları puanlar arasındaki tutarlılık incelenir. Madde sayısı çok, konular basitten zora yerleştirilmiş veya kümelenmişse bu yöntem uygun değildir.	Tek konuyu değerlendiren testler
Kuder- Richardson 20 ve 21	KR20 ve KR21 olarak kısaltılan bu yöntemle testin iç tutarlılığı değerlendirilir.	0-1 puanlı ya da evet-hayır seçenekli testler gibi
Cronbach Alfa	Ölçme aracının puanlaması çok kategorili olduğunda uygulanır.	Çok kategorili puanlama testleri

TEST GELİŞTİRME VE MADDE (SORU) TÜRLERİ

Testi Alan Kişi Sayısına Göre	Testin Uygulanış Süresine Göre	Testin Ölçtüğü Niteliğe Göre	Değerlendirme Yaklaşımına Göre	Hazırlanış Biçimine Göre	Veri Toplama Tekniğine Göre
1. Bireysel	1. Süreli	1. Hız testleri	1. Objektif	1. Standart	1.Performans
2. Grup	2. Süresiz	2. Güç Testleri	2. Subjektif	2. Öğretmen yapımı	2. Kağıt-Kalem

Test Geliştirme İşlem Basamakları

Standart testler için:

- → Amacın belirlenmesi
- → Kapsamın belirlenmesi, belirtke tablosunun hazırlanması
- → Denemelik maddelerin yazılması
- → Redaksiyon
- → Denemelik test formunun hazırlanması
- → Testin uygulanması
- → Test madde istatistiklerinin hesaplanması
- → Seçilen maddelerden oluşan nihai formun oluşturulması

Sınıf içi testler için:

- ✓ Amaç, davranış, konuların belirlenmesi
- ✓ Sınav günü ve süresinin belirlenmesi
- ✓ Belirtke tablosunun hazırlanması
- ✓ Madde türü ve sayısının belirlenmesi
- ✓ Soru güçlük dağılımının belirlenmesi
- ✓ Yöntemlerin belirlenmesi
- ✓ Cevap anahtarı ve puanlamanın oluşturulması
- ✓ Yazma ve çoğaltma
- ✓ Yönerge hazırlanması
- ✓ Uygulama
- ✓ Test ve madde istatistiklerinin hesaplanması



- ✓ Doğru- yanlış
- ✓ Eşleştirme
- ✓ Çoktan seçmeli

→ Sistematik hata oranı düşük, puanlama objektiftir. Şans olasılığı vardır.

<u>Açık uçlu soru tipi</u>

- ✓ Yazılı yoklama
- ✓ Boşluk doldurma
- ✓ Kısa yanıtlı
- ✓ Sözlü yoklama

→ Cevaplar sınırsız özgürlük içerir. Hazırlaması kolay ancak değerlendirmesi zordur. Kısmen sübjektiftir.

DESTEKLEYİCİ DEĞERLENDİRME YAKLAŞIMLARI

- PORTFOLYO

*Süreci yansıtan ve ürünü yansıtan olmak üzere iki türü vardır. İçeriği,

- Öğretmen kayıtları
- Öğrenci çalışmaları
- Günlükler
- Aileye veya öğrenciye yazılan mektuplar
- Video, ses kayıtları

*Elektronik ortamda tutulana e-portfolyo denir.
*Toplama-seçme- yansıtma- sonuç süreciyle oluşturulur.

- PERFORMANS DEĞERLENDİRME

Öğrencinin yaptıkları: Öz değerlendirme, akran değerlendirme
Öğretmenin yaptıkları: Rubrik (Dereceli puanlama anahtarı), kontrol listesi, derecelendirme ölçeği, gözlem formu

TEST İSTATİSTİKLERİ: Ölçme işlemi sonucunda bireylerin toplam puanları üzerinden hesaplanan istatistiklerdir. Eğitimde en sık kullanılan istatistikler merkezî eğilim ölçüleri, değişkenlik ölçüleri, dağılım özellikleri, standart puanlar vb.dir.
Merkezî eğilim ölçüleri: aritmetik ortalama, medyan (ortanca) ve moddur. Bu üç istatistiğin de temel işlevi bir puan dizisindeki merkezi bulmaktır. Değişkenlik ölçüleri ranj (dizi genişliği), varyans, standart sapma ve çeyrek sapmadır.

MADDE İSTATİSTİKLERİ: Ölçme işlemi sonucunda bireylerin madde puanları üzerinden hesaplanan istatistiklerdir. Eğitimde en sık kullanılan istatistikler madde güçlükindeksi, madde ayırt edicilik indeksi, madde güvenilirliği vb.dir.

Madde güçlük indeksi; maddenin kabaca zor mu, orta mı, kolay mı olduğunu gösteren bir istatistiktir.

EĞİTİM İZLEME ARAŞTIRMALARI

1-Uluslararası İzleme Çalışmaları

ARAŞTIRMA ADI VE İÇERİĞİ	PISA	PIRLS	TIMSS
	Matematik+fen+okuma becerileri	Okuma becerileri	Matematik+fen
KİM DÜZENLER?	OECD	IEA	IEA
SÜRE ARALIĞI	3 yılda bir	5 yılda bir	4 yılda bir
YAŞ GRUBU	15	10	4 ve 8. sınıf

- PISA
  1- 2000- okuma becerisi, 2003- matematik okuryazarlığı, 2006 fen okur yazarlığı şeklinde sırasıyla çapraz modelleme uygulanır.
  2- Ülkemizdeki örneklem seçimi ülkenin sosyo-ekonomik farklılıklarını dikkate alan 12 bölge üzerinden yapılmaktadır.
  3- Türkiye, 2003 yılından beri katılmaktadır. 79 ülke katılımı vardır.
  4- Hedef OECD (Ekonomik Kalkınma ve İş Birliği Örgütü) ortalamasına ulaşmak ve geçmektir.

- TIMMS
  1- Matematik ve fen okuryazarlığını programla ilişkilendirerek ölçer.
  2- Uluslararası Eğitim Başarıları Değerlendirme Kuruluşu (IEA) tarafından ilk defa 1995, ülkemizde ilk defa 1999 yılında uygulanmıştır.
  3- Yedinci döngüsünde 64 ülke yer almıştır.
  4- Ülkemizde 12 bölgeden tesadüfi örneklemle uygulanır.
  5- Türkiye'de diğer ülkelerdeki gibi anasınıfı zorunlu olmadığından 2019 yılından beri 4. Sınıf değil 5. Sınıf öğrencileri sınava katılmaktadır.
  6- Türkiye TIMSS'de PISA'ya göre daha başarılıdır, çünkü TIMSS eğitim programına dayalı ölçme yapar.

- PIRLS
  Türkiye 2001 ve 2021 yıllarında iki defa katılmıştır.

- OECD Sosyal ve Duygusal Beceriler Araştırması
  1- 2017 yılında başlamıştır.
  2- Türkiye sadece İstanbul'dan seçilen örneklemlerle katılmıştır.
  3- Ölçümler beş faktörlü kurama (Big five) göre yapılmaktadır.
  4- Veli-öğretmen- öğrenci veri çeşitlemesiyle veriler toplanır.
  5- Kanada, ABD, Türkiye, Kolombiya, Portekiz, Finlandiya, Rusya, Çin, Güney Kore katılmıştır.
  6- 10 ve 15 yaş gruplarıyla çalışılmıştır.
  7- Okula aidiyet, sanat etkinlikleri, yaratıcılık, okul iklimi, empati düzeyi vb. durumlar ölçülerek sosyal ve duygusal becerilere dair veriler elde edilmektedir.

- ABİDE
1- Akademik Becerilerin İzlenmesi ve Değerlendirilmesi araştırmasıdır.
2- Ulusaldır.
3- Üst düzey bilişsel becerileri ölçer.

- TMF-ÖBA
1- Türkçe, Matematik, Fen Öğrenci Başarıları İzleme Araştırması'dır.
2- Eğitim programına bağlı kazanımlara yönelik inceleme yapar.